

**Sullivan, Daniel**

---

**From:** Sullivan, Daniel  
**Sent:** Tuesday, September 27, 2005 7:56 AM  
**To:** STIC-Biotech/ChemLib  
**Subject:** Sequence search 09/982091

Please search for the following in the pending, issued patent and commercial databases:

A nucleic acid encoding SEQ ID NO: 4;  
A nucleic acid encoding SEQ ID NO: 2;  
A nucleic acid comprising residues 1-500 of SEQ ID NO: 1;  
A nucleic acid comprising residues 4200-4740 of SEQ ID NO: 1;  
A nucleic acid comprising residues 1-500 of SEQ ID NO: 5;  
A nucleic acid comprising residues 58300-58837 of SEQ ID NO: 5

Thank you.

---

**Daniel M. Sullivan**

Examiner AU 1636  
Remsen Bldg.  
Room 2A74

Tel: (571) 272-0779

Mailbox: 2C70

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 21:21:46 ; Search time 1847.42 Seconds  
(without alignments)  
1410.995 Million cell updates/sec

Title: US-09-982-091A-5\_COPY\_58300\_58837

Perfect score: 538

Sequence: 1 caggtacagcaggagagcga.....tatgccagcagtcctcctgag 538

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenBml:\*

1: gb\_ba:\*\n2: gb\_htg:\*\n3: gb\_in:\*\n4: gb\_om:\*\n5: gb\_ov:\*\n6: gb\_pat:\*\n7: gb\_ph:\*\n8: gb\_pl:\*\n9: gb\_pr:\*\n10: gb\_ro:\*\n11: gb\_sts:\*\n12: gb\_sy:\*\n13: gb\_un:\*\n14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	538	100.0	194296	9 AL354864	AL354864 Human DNA
2	361	67.1	2567	9 BC026909	BC026909 Homo sapi
3	361	67.1	2615	6 AR338531	AR338531 Sequence
4	361	67.1	2934	6 AX834078	AX834078 Sequence
5	361	67.1	2934	9 AK096303	AK096303 Homo sapi
6	356	66.2	1046	6 CQ727306	CQ727306 Sequence
7	300.2	55.8	2394	6 AX834179	AX834179 Sequence
8	300.2	55.8	2394	9 AK096466	AK096466 Homo sapi
9	255.2	47.4	213391	10 AL606935	AL606935 Mouse DNA
10	242	45.0	223363	2 AC120700	AC120700 Rattus no
11	242	45.0	247462	2 AC134478	AC134478 Rattus no
12	227	42.2	203980	2 AC134058	AC134058 Rattus no
13	51.6	9.6	125020	9 AF429315	AF429315 Homo sapi
14	43	8.0	125020	9 AF429315	AF429315 Homo sapi
15	39	7.2	364	6 AR425705	AR425705 Sequence
16	39	7.2	364	6 AX986399	AX986399 Sequence
17	39	7.2	364	6 BD121258	BD121258 EST and e
18	38.2	7.1	210813	10 AL627070	AL627070 Mouse DNA
19	37.8	7.0	135851	9 AC091714	AC091714 Papio anu

20	37.6	7.0	212050	1 AL646060	AL646060 Ralstonia
21	37.4	7.0	415	6 CQ451255	CQ451255 Sequence
22	37.4	7.0	1605	6 BD180291	BD180291 Highly th
23	37.4	7.0	25593	3 CEH21P03	295619 Caenorhabdi
24	37.4	7.0	144734	2 AC148922	AC148922 Dasytus n
25	37.4	7.0	179439	2 AC131897	AC131897 Oryctolag
26	37.4	7.0	211730	10 AC132621	AC132621 Mus muscu
27	37.4	7.0	252686	2 AC106177	AC106177 Rattus no
28	37.4	7.0	265229	2 AC121468	AC121468 Rattus no
29	37.4	7.0	274676	1 AE017305	AE017305 Thermus t
30	37.2	6.9	3120	6 CQ717965	CQ717965 Sequence
31	37.2	6.9	3120	9 HSU58917	U58917 Homo sapien
32	37.2	6.9	3223	6 AR034066	AR034066 Sequence
33	37.2	6.9	3223	6 AR097733	AR097733 Sequence
34	37.2	6.9	3223	6 AR097733	AR097733 Sequence
35	37.2	6.9	3223	6 AR102359	AR102359 Sequence
36	37.2	6.9	3223	6 AR104870	AR104870 Sequence
37	37.2	6.9	3223	6 AR131081	AR131081 Sequence
38	37.2	6.9	3223	6 AR137603	AR137603 Sequence
39	37.2	6.9	3223	6 AR453137	AR453137 Sequence
40	37.2	6.9	3223	6 BD023246	BD023246 Method fo
41	37.2	6.9	3373	9 BC011624	BC011624 Homo sapi
42	37.2	6.9	51574	2 AC138651	AC138651 Homo sapi
43	37.2	6.9	82806	9 AC007111	AC007111 Homo sapi
44	37.2	6.9	101269	9 AC005300	AC005300 Homo sapi
45	37.2	6.9	149061	9 AC098805	AC098805 Homo sapi

ALIGNMENTS

RESULT 1  
AL354864  
LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL354864 194296 bp DNA linear PRI 04-APR-2001  
Human DNA sequence from clone RP11-435D7 on chromosome 1. Contains  
ESTs, STSs, GSSs and Cpg islands. Contains a novel gene and the 5'  
part of the PSMB2 gene for proteasome (prosome, macropain) subunit  
2, beta type, complete sequence.  
AL354864 GI:10185566  
AL354864.16 GI:10185566  
HTG; Cpg island; macropain; prosome; proteasome; PSMB2.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 194296)  
Howden, P.  
Direct Submission  
Submitted (06-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk  
On Sep 19, 2000 this sequence version replaced gi:9714820.  
During sequence assembly data is compared from overlapping clones.  
where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping  
Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr1  
This sequence is the entire insert of clone RP11-435D7 The true  
left end of clone RP1-62B3 is at 123829 in this sequence. The true  
right end of clone RP4-555P23 is at 81044 in this sequence. This  
sequence has been finished according to sequence map criteria as  
follows. An attempt is made to resolve all sequencing problems,  
such as compressions and repeats, but not necessarily within known

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 21:21:46 ; Search time 1716.93 Seconds  
(without alignments)  
1410.995 Million cell updates/sec

Title: US-09-982-091A-5\_COPY\_1\_500

Perfect score: 500

Sequence: 1 aagcaggtagcttctaactt.....tgccaccagcgctgagtgcc 500

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_cm:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	500	100.0	194296	9	AL354864	AL354864 Human DNA
2	498.4	99.7	111862	2	AL139143	AL139143 Homo sapi
3	61.8	12.4	87307	9	AC135791	AC135791 Homo sapi
4	61.8	12.4	132899	9	AC093506	AC093506 Homo sapi
5	61.8	12.4	153225	2	AC013581	AC013581 Homo sapi
6	61.8	12.4	153538	9	AC093279	AC093279 Homo sapi
7	61.8	12.4	174039	9	AC138853	AC138853 Homo sapi
8	61.8	12.4	186972	9	AC137794	AC137794 Homo sapi
9	61.8	12.4	189670	2	AC041047	AC041047 Homo sapi
10	61.8	12.4	190684	2	AC139798	AC139798 Homo sapi
11	61.8	12.4	193666	2	AC140124	AC140124 Homo sapi
12	61.8	12.4	222307	2	AC140514	AC140514 Homo sapi
13	60.4	12.1	151991	2	AC053535	AC053535 Homo sapi
14	60.4	12.1	159112	9	AC040975	AC040975 Homo sapi
15	59.8	12.0	155942	2	AC136691	AC136691 Homo sapi
16	59.8	12.0	161765	9	AC113190	AC113190 Homo sapi
17	59.8	12.0	176479	2	AC135631	AC135631 Homo sapi
18	59.8	12.0	337344	6	CQ869869	CQ869869 Sequence
19	59.2	11.8	144319	9	AC112218	AC112218 Homo sapi

20	59.2	11.8	162997	9	AC092035	AC092035 Homo sapi
21	58.8	11.8	172660	9	AC136968	AC136968 Pan trogl
22	58.8	11.8	255818	2	AC145871	AC145871 Pan trogl
23	58.4	11.7	41924	2	AC087403	AC087403 Homo sapi
24	58.4	11.7	165459	9	AC090587	AC090587 Homo sapi
25	58.4	11.7	173720	9	AC097264	AC097264 Pan trogl
26	58.4	11.7	181835	9	AC091152	AC091152 Homo sapi
27	58.4	11.7	184106	2	AC060812	AC060812 Homo sapi
28	58	11.6	149751	2	AC048339	AC048339 Homo sapi
29	58	11.6	176999	2	AC053518	AC053518 Homo sapi
30	58	11.6	181445	9	AC026367	AC026367 Homo sapi
31	57.8	11.6	67372	9	AL138741	AL138741 Human DNA
32	57.2	11.4	143981	9	HSJ1050K3	AL121776 Human DNA
33	57	11.4	108687	9	AL356320	AL356320 Human DNA
34	57	11.4	172837	9	AL450304	AL450304 Human DNA
35	56.8	11.4	119063	9	AC027311	AC027311 Homo sapi
36	56.8	11.4	172239	2	AC079307	AC079307 Homo sapi
37	56.6	11.3	628	6	BD275472	BD275472 47 Human
38	56.6	11.3	80149	2	AC025995	AC025995 Homo sapi
39	56.6	11.3	91654	9	AL357118	AL357118 Human DNA
40	56.6	11.3	130027	9	AC004982	AC004982 Homo sapi
41	56.6	11.3	148852	9	AC005878	AC005878 citb_255
42	56.6	11.3	163652	2	AC145845	AC145845 Pan trogl
43	56.6	11.3	165002	2	AL355814	AL355814 Homo sapi
44	56.6	11.3	175020	9	AC024267	AC024267 Homo sapi
45	56.6	11.3	177773	9	AC010761	AC010761 Homo sapi

ALIGNMENTS

RESULT 1	AL354864	194296 bp	DNA	linear	PRI 04-APR-2001
LOCUS	AL354864				
DEFINITION	Human DNA sequence from clone RPl1-435D7 on chromosome 1. Contains ESTs, STSs, GSSs and Cpg islands. Contains a novel gene and the 5' part of the PSMB2 gene for proteasome (prosome, macropain) subunit 2, beta type, complete sequence.				
ACCESSION	AL354864				
VERSION	AL354864.16	GI:10185566			
KEYWORDS	HTG; Cpg island; macropain; prosome; proteasome; PSMB2.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	1 (bases 1 to 194296)				
TITLE	Howden, P.				
JOURNAL	Direct Submission				
COMMENT	Submitted (06-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk requests: clonerequest@sanger.ac.uk On Sep 19, 2000 this sequence version replaced gi:9714820. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr1 This sequence is the entire insert of clone RPl1-435D7 The true left end of clone RPl1-62B3 is at 123829 in this sequence. The true right end of clone RPl1-555P23 is at 81044 in this sequence. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known				



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2005, 00:01:51 ; Search time 89.0197 Seconds  
(without alignments)  
9889.010 Million cell updates/sec

Title: US-09-982-091A-5\_COPY\_58300\_58837

Perfect score: 538  
Sequence: 1 caggtacagcaggagagcaga.....tatgccagcagttcctctgag 538

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	361	67.1	2615	4 US-09-620-312D-22	Sequence 22, Appl
2	39	7.2	364	4 US-09-621-976-17202	Sequence 17202, A
3	37.2	6.9	3223	2 US-08-620-694A-9	Sequence 9, Appli
4	37.2	6.9	3223	3 US-09-022-255-9	Sequence 9, Appli
5	37.2	6.9	3223	3 US-09-022-696-9	Sequence 9, Appli
6	37.2	6.9	3223	3 US-08-978-773-3	Sequence 3, Appli
7	37.2	6.9	3223	3 US-09-022-253-9	Sequence 9, Appli
8	37.2	6.9	3223	3 US-09-022-260-9	Sequence 9, Appli
9	37.2	6.9	3223	3 US-09-022-259-9	Sequence 9, Appli
10	37.2	6.9	3223	3 US-09-022-257-9	Sequence 9, Appli
11	37.2	6.9	3223	4 US-09-549-679-9	Sequence 9, Appli
12	37.2	6.9	16154	4 US-09-949-016-16669	Sequence 11, Appl
13	36.4	6.8	8916	3 US-09-579-181-11	Sequence 10, Appl
14	36.4	6.8	9354	3 US-09-579-181-10	Sequence 11, Appl
15	35	6.5	3154	4 US-09-949-016-2101	Sequence 2101, Ap
16	34.8	6.5	63319	4 US-09-949-016-16579	Sequence 16579, A
17	34	6.3	771	2 US-08-972-008-3	Sequence 3, Appli
18	34	6.3	771	2 US-08-972-409-3	Sequence 3, Appli
19	34	6.3	1349	4 US-09-949-016-1391	Sequence 1391, Ap
20	34	6.3	1740	3 US-08-796-101-50	Sequence 50, Appl
21	34	6.3	2495	3 US-09-141-027-1	Sequence 1, Appli
22	34	6.3	2495	4 US-09-617-804-1	Sequence 1, Appli
23	34	6.3	2500	4 US-09-949-016-791	Sequence 791, App
24	34	6.3	2525	2 US-08-972-008-1	Sequence 1, Appli
25	34	6.3	2525	3 US-09-267-409-1	Sequence 1, Appli
26	33.8	6.3	1127	4 US-09-620-312D-254	Sequence 254, App
27	33.8	6.3	8310	3 US-08-870-126-11	Sequence 11, Appl

28	33.8	6.3	8310	3 US-09-445-247-11	Sequence 11, Appl
29	33.8	6.3	14985	1 US-08-652-972A-6	Sequence 6, Appli
30	33.8	6.3	14985	5 PCT-US96-06231A-6	Sequence 6, Appli
31	33.8	6.3	21856	4 US-09-949-016-12205	Sequence 12205, A
32	33.4	6.2	2356	4 US-09-799-451-18	Sequence 18, Appl
33	33.4	6.2	32010	4 US-09-949-016-13127	Sequence 13127, A
34	33.4	6.2	1230025	4 US-09-198-452A-1	Sequence 1, Appli
35	33.4	6.2	1230230	4 US-09-438-185A-1	Sequence 1, Appli
36	33.2	6.2	2549	3 US-09-467-082-3	Sequence 3, Appli
37	33.2	6.2	2549	4 US-09-394-455-5	Sequence 5, Appli
38	33.2	6.2	2549	4 US-09-949-016-3584	Sequence 3584, Ap
39	33.2	6.2	2608	4 US-09-394-455-35	Sequence 35, Appl
40	33.2	6.2	12425	4 US-09-616-289-50	Sequence 50, Appl
41	33.2	6.2	29930	4 US-09-949-016-15326	Sequence 15326, A
42	33	6.1	2831	2 US-08-808-982-3	Sequence 3, Appli
43	33	6.1	2831	3 US-09-306-902A-3	Sequence 3, Appli
44	32.8	6.1	981	4 US-09-252-991A-13623	Sequence 13623, A
45	32.8	6.1	1056	4 US-09-252-991A-13421	Sequence 13421, A

ALIGNMENTS

RESULT 1  
US-09-620-312D-22  
; Sequence 22, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yungding  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_fl\_genes Version 1.0  
; SEQ ID NO. 22  
; LENGTH: 2615  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (160)..(849)  
; US-09-620-312D-22

Query Match 67.1%; Score 361; DB 4; Length 2615;  
Best Local Similarity 100.0%; Pred. No. 2.2e-96;  
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 GCTAGATGTTGCGCATCCAGCCAGGCTAGCTGAGGGGGCCAAATTCCTGGGGAGCCAC 237  
Db 155 GCTAGATGTTGCGCATCCAGCCAGGCTAGCTGAGGGGGCCAAATTCCTGGGGAGCCAC 214

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 21:10:56 ; Search time 278.963 seconds  
(without alignments)  
11416.644 Million cell updates/sec

Title: US-09-982-091A-5\_COPY\_58300\_58837

Perfect score: 538  
Sequence: 1 caggtacagcagggagagcga.....tatgccagcagtcctcctgag 538

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: geneseqn19808:\*  
2: geneseqn19908:\*  
3: geneseqn20008:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	538	100.0	58837	6	ABK52612	Abk52612 Human Cla
2	361	67.1	2615	4	AA158146	Aa158146 Human pol
3	361	67.1	2615	5	ADQ98352	Adq98352 DNA encod
4	361	67.1	2615	9	ADB48112	Adb48112 Novel hum
5	361	67.1	2934	11	ADM02517	Adm02517 Human CDN
6	359.4	66.8	2616	4	AA159932	Aa159932 Human pol
7	300.2	55.8	2394	11	ADM02618	Adm02618 Human CDN
8	38	7.1	10902	5	AA590570	Aa590570 DNA encod
9	37.4	7.0	415	6	ABN24269	Abn24269 Human ORF
10	37.2	6.9	3120	10	ADD25545	Add25545 Binding d
11	37.2	6.9	3120	10	ADK61414	Adk61414 Ovarian c
12	37.2	6.9	3120	12	ADP09696	Adp09696 Human IL1
13	37.2	6.9	3120	12	ADP13348	Adp13348 Renal cel
14	37.2	6.9	3223	2	AA133801	Aa133801 Human int
15	37.2	6.9	3223	2	AAV27592	Aav27592 Human int
16	37.2	6.9	3223	2	AA101922	Aa101922 Human IL-
17	37.2	6.9	3223	3	AA51988	Aa51988 Human int
18	37.2	6.9	3223	3	AA52146	Aa52146 Human int
19	37.2	6.9	3223	3	AA59871	Aa59871 Human int
20	37.2	6.9	3223	3	AA61240	Aa61240 Human IL-

21	37.2	6.9	3223	4	AA57188	Aa57188 Human IL-
22	37.2	6.9	3223	4	AA02815	Aa02815 Human int
23	37.2	6.9	3223	12	ADJ88264	Adj88264 Human IL-
24	37.2	6.9	3223	12	ADL24268	Adl24268 Human IL-
25	37.2	6.9	3223	12	ADQ91251	Adq91251 Human CDN
26	37.2	6.9	3223	12	ADR01307	Adr01307 Human int
27	37.2	6.9	3429	10	ADG32974	Adg32974 Human DNA
28	36.4	6.8	2000	8	ADA71938	Ada71938 Rice gene
29	36.4	6.8	2244	5	AA590569	Aa590569 DNA encod
30	36.4	6.8	5980	5	ABA82988	Ab82988 Human tra
31	36.4	6.8	7666	13	ADR07802	Adr07802 Full leng
32	36.4	6.8	8916	5	AAC89860	Aac89860 Human SRC
33	36.4	6.8	9126	10	ADC35126	Adc35126 Human bre
34	36.4	6.8	9354	5	AAC89859	Aac89859 Human SRC
35	36.4	6.8	9958	8	ACC46296	Acc46296 Human dit
36	36.4	6.8	10795	3	AAC75440	Aac75440 Human ORF
37	36	6.7	5349	3	AAA65527	Aaa65527 Porcine B
38	35.8	6.7	2895	12	ADO10076	Ado10076 Novel hum
39	35.4	6.6	1479	12	ADJ40245	Adj40245 Plant CDN
40	35	6.5	34981	10	ADC87619	Adc87619 Human GPC
41	34.8	6.5	20015	5	ABA20740	Ab20740 Human ner
42	34.6	6.4	4348	5	AA545069	Aa545069 cDNA enco
43	34.6	6.4	4401	10	ADF92402	Adf92402 Human ubi
44	34.6	6.4	4474	3	AAC76399	Aac76399 Human ORF
45	34.6	6.4	4506	12	ADL14063	Adl14063 Human CDN

ALIGNMENTS

RESULT 1  
ABK52612 ID ABK52612 standard; DNA; 58837 BP.  
XX ABK52612;  
AC  
XX 27-AUG-2002 (first entry)  
DT  
XX Human Claspin genomic sequence.  
DE  
XX Chk1 protein; SQ/TQ motif; isoelectric point; cell cycle progression; ds;  
KW nuclear localisation signal; DNA replication checkpoint; benign neoplasm;  
KW cell proliferative disorder; malignant neoplasm; human; claspin.  
XX  
OS Homo sapiens.  
XX  
XX WO200233115-A2.  
XX  
XX 25-APR-2002.  
PD  
XX  
XX 17-OCT-2001; 2001WO-US032316.  
PF  
XX 17-OCT-2000; 2000US-0241246P.  
PR  
XX (CALY ) CALIFORNIA INST OF TECHNOLOGY.  
XX  
XX Kumagai A, Dunphy WG;  
XX WPI, 2002-454610/48.  
DR  
XX  
XX Novel Claspin polypeptide specifically interacting with chk1 protein  
PT useful for identifying compound that modulates cell cycle progression and  
PT for treating cell proliferative disorder like neoplasm.  
XX  
XX Claim 10; Fig 7; 97pp; English.  
XX  
XX The present invention relates to a new substantially pure Claspin  
CC polypeptide that specifically interacts with a Chk1 protein, having SQ/TQ  
CC motifs, an isoelectric point of 4.5 and at least one nuclear localisation  
CC signal. The method of the invention is useful for identifying a compound  
CC that modulates cell cycle progression and for modulating cell cycle  
CC progression in a cell. The invention is useful for the proper operation  
CC of DNA replication checkpoint in the cell cycle. The method is also



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 23:18:16 ; Search time 1770.04 Seconds  
(without alignments)  
11569.535 Million cell updates/sec

Title: US-09-982-091A-5\_COPY\_58300\_58837

Perfect score: 538  
Sequence: 1 caggtagacaggaggagacga.....tatgccagcagttccctgag 538

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_hc: \*  
4: gb\_est3: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_g881: \*  
9: gb\_g882: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	438.8	81.6	987	4	BM562991	BM562991 AGENCOURT
2	361	67.1	667	4	BG704820	BG704820 602688458
3	361	67.1	719	1	AL042728	AL042728 DKFZp434B
4	361	67.1	848	4	BI916179	BI916179 603178063
5	361	67.1	872	5	BU158384	BU158384 AGENCOURT
6	361	67.1	879	5	BU189176	BU189176 AGENCOURT
7	361	67.1	897	5	BQ878792	BQ878792 AGENCOURT
8	361	67.1	1047	4	BM544506	BM544506 AGENCOURT
9	361	67.1	1060	5	BM925770	BM925770 AGENCOURT
10	361	67.1	1111	5	BM926654	BM926654 AGENCOURT
11	361	67.1	1124	5	BM926654	BM926654 AGENCOURT
12	361	67.1	2539	3	CR621322	CR621322 full-leng
13	359.4	66.8	582	5	BP228770	BP228770 BP228770
14	356	66.2	612	7	CV029336	CV029336 8072 Full
15	356	66.2	690	9	AY417699	AY417699 Homo sapi
16	353	65.6	874	6	CB993658	CB993658 AGENCOURT
17	349	64.9	730	4	BI522572	BI522572 603175275
18	349	64.9	775	4	BI768300	BI768300 603056494
19	349	64.9	804	4	BG613708	BG613708 602639876
20	347.8	64.6	690	9	AY417700	AY417700 Pan trogl
21	341.8	63.5	956	5	BX460530	BX460530 BX460530
22	341.8	63.5	967	1	AL536618	AL536618 AL536618
23	340.2	63.2	603	4	BI823682	BI823682 603038591
24	339	63.0	848	4	BG819863	BG819863 602781937

25	323.6	60.1	501	6	CA391404	CA391404 cs14h09.Y
26	285.4	53.0	725	7	CN294501	CN294501 170006001
27	282.4	52.5	854	2	BE786197	BE786197 601474318
28	272	50.6	566	1	AV589245	AV589245 AV589245
29	242.8	45.1	668	2	BB618564	BB618564 BB618564
30	242.8	45.1	693	9	AY417701	AY417701 Mus muscu
31	242.8	45.1	1947	3	AK077438	AK077438 Mus muscu
32	218	40.5	926	4	BI912192	BI912192 B63 AFT02
33	217.8	40.5	583	6	CB586472	CB586472 AMGNNUC:N
34	215.2	40.0	583	6	CB586472	CB586472 AMGNNUC:N
35	210	39.0	897	5	BQ433497	BQ433497 AGENCOURT
36	196.4	36.5	694	4	BI695741	BI695741 603345090
37	194.4	36.1	421	1	AA084945	AA084945 zn13a11.r
38	166.4	30.9	390	2	BF892134	BF892134 QV1-MT013
39	156.8	29.1	168	7	CR541130	CR541130 DKFZp4591
40	142	26.4	212	7	T05523	T05523 EST03412 Fe
41	138.4	25.7	403	6	CB768787	CB768787 AMGNNUC:M
42	133	24.7	496	4	BI464515	BI464515 603206108
43	118.2	22.0	336	4	BI134471	BI134471 UI-M-BH3-
44	116.6	21.7	454	6	CB784968	CB784968 AMGNNUC:N
45	115	21.4	1183	4	BM472927	BM472927 AGENCOURT

ALIGNMENTS

RESULT 1  
BM562991  
LOCUS  
DEFINITION  
AGENCOURT\_6566840 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:5736829  
5', mRNA sequence.  
BM562991  
BM562991.1 GI:18809535  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 987)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM12745 row: p column: 14  
High quality sequence stop: 631.  
Location/Qualifiers  
1. . 987  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5736829"  
/tissue\_type="duodenal adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_88"  
/note="Organ: small intestine; Vector: pCMV-SPORT6;  
Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally;  
oligo-dT primed. Average insert size 1.767 kb. Library  
enriched for full-length clones and constructed by Life  
Technologies. Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 81.6%; Score 438.8; DB 4; Length 987;  
Best Local Similarity 93.9%; Pred. No. 1.3e-108;  
Matches 489; Conservative 0; Mismatches 28; Indels 4; Gaps 3;

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 21:10:56 ; Search time 259.259 Seconds  
(without alignments)  
11416.644 Million cell updates/sec

Title: US-09-982-091a-5\_COPY\_1\_500  
Perfect score: 500  
Sequence: 1 aagcagtagtcttaactt.....tgccaccagcgtagtgac 500

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: geneseqn1980a:\*

2: geneseqn1990a:\*

3: geneseqn2000a:\*

4: geneseqn2001a:\*

5: geneseqn2001b:\*

6: geneseqn2002a:\*

7: geneseqn2002b:\*

8: geneseqn2003a:\*

9: geneseqn2003b:\*

10: geneseqn2003c:\*

11: geneseqn2003d:\*

12: geneseqn2004a:\*

13: geneseqn2004b:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	500	100.0	58837	6	ABK52612	Abk52612 Human Cla
2	59.8	12.0	337344	13	ABD32715	Abd32715 Human can
3	56.6	11.3	628	3	AAC69419	Aac69419 Human sec
4	56.2	11.2	43981	10	ADC86790	Adc86790 Human GPC
5	56.2	11.2	50807	12	ADQ97804	Adq97804 Human can
6	56	11.2	9219	6	ABK83482	Abk83482 Human CDN
7	55.8	11.2	129722	6	ABQ88117	Abq88117 Human oes
8	55.8	11.2	129722	12	ADQ18027	Adq18027 Human sof
9	55.4	11.1	670	3	AAA81663	Aaa81663 N. mening
10	55.4	11.1	57038	11	ADP65343	Adp65343 Human seq
11	55.4	11.1	92794	11	ADP65471	Adp65471 Human seq
12	55.2	11.0	73063	12	ADQ97727	Adq97727 Human can
13	54.8	11.0	27007	4	AAS27764	Aas27764 DNA encod
14	54.8	11.0	27007	10	ADB94567	Adb94567 Novel hum
15	54.8	11.0	90400	12	ADP08389	Adp08389 Human chr
16	54.8	11.0	105499	13	ABD33280	Abd33280 Human can
17	54.6	10.9	110000	12	ADQ97960	Adq97960 Human can
18	54.4	10.9	362	6	ABO59792	Abq59792 Human col
19	54.4	10.9	4350	13	ADR08235	Adr08235 Full leng
20	54.2	10.8	71887	11	ACN44950	Acn44950 Human gen

C	21	54	10.8	176001	12	ADK43203	Adk43203 Human pro
C	22	54	10.8	186739	12	ADK43195	Adk43195 Human pro
C	23	54	10.8	220224	11	ACN44702	Acn44702 Human gen
C	24	53.8	10.8	2656	4	AAH17941	Aah17941 Human CDN
C	25	53.8	10.8	12312	5	AAF97872	Aaf97872 Human neu
	26	53.6	10.7	347814	12	ADQ59440	Adq59440 Human can
	27	53.2	10.6	110000	5	AAF84800	Aaf84800 Nucleotid
	28	53.2	10.6	191150	12	ADM69029	Adm69029 Human pla
	29	53.2	10.6	245531	13	ABD33022	Abd33022 Human can
	30	53	10.6	558	5	ABV38118	Abv38118 Human pro
	31	53	10.6	1359	8	ABZ74550	Abz74550 Secreted
	32	53	10.6	1359	10	ABZ68084	Abz68084 Human sec
	33	53	10.6	2565	5	ABV25262	Abv25262 Human pro
C	34	53	10.6	3089	4	AAK72379	Aak72379 Human imm
	35	53	10.6	18902	8	ABZ73865	Abz73865 Secreted
	36	53	10.6	18902	10	ABZ67444	Abz67444 Human sec
	37	53	10.6	110000	6	ABA90193	Ab90193 Human oes
	38	53	10.6	110000	6	ABQ87681	Abq87681 Human oes
	39	53	10.6	110000	8	ABX33717	Abx33717 Gene enco
	40	53	10.6	168174	6	ABT11173	Abt11173 Human 5-1
	41	53	10.6	168273	6	ABT11114	Abt11114 Human 5-1
	42	52.8	10.6	6136	12	ADQ64370	Adq64370 Novel hum
	43	52.8	10.6	53242	11	ACN43926	Acn43926 Human gen
C	44	52.8	10.6	68940	2	AAK57351	Aak57351 Human chr
	45	52.8	10.6	165156	13	ADS36459	Ads36459 Human aut

ALIGNMENTS

RESULT 1	ABK52612	standard; DNA; 58837 BP.
ID	ABK52612	
XX	ABK52612;	
AC	27-AUG-2002	(first entry)
DT		
XX	Human Claspin genomic sequence.	
DE		
XX		
KW	Chk1 protein; SQ/TQ motif; isoelectric point; cell cycle progression; ds;	
KW	nuclear localisation signal; DNA replication checkpoint; benign neoplasm;	
KW	cell proliferative disorder; malignant neoplasm; human; claspin.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200233115-A2.	
XX		
PD	25-APR-2002.	
XX		
PF	17-OCT-2001; 2001WO-US032316.	
XX		
PR	17-OCT-2000; 2000US-0241246P.	
XX		
PA	(CALY ) CALIFORNIA INST OF TECHNOLOGY.	
XX		
PI	Kumagai A, Dunphy WG;	
XX		
DR	WPI; 2002-454610/48.	
XX		
PT	Novel Claspin polypeptide specifically interacting with chk1 protein	
PT	useful for identifying compound that modulates cell cycle progression and	
PT	for treating cell proliferative disorder like neoplasm.	
XX		
PS	Claim 10; Fig 7; 97pp; English.	
XX		
CC	The present invention relates to a new substantially pure Claspin	
CC	polypeptide that specifically interacts with a Chk1 protein, having SQ/TQ	
CC	motifs, an isoelectric point of 4.5 and at least one nuclear localisation	
CC	signal. The method of the invention is useful for identifying a compound	
CC	that modulates cell cycle progression and for modulating cell cycle	
CC	progression in a cell. The invention is useful for the proper operation	
CC	of DNA replication checkpoint in the cell cycle. The method is also	



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2005, 02:35:27 ; Search time 687.831 Seconds  
(without alignments)  
5019.154 Million cell updates/sec

Title: US-09-982-091A-5\_COPY\_1\_500  
Perfect score: 500  
Sequence: 1 aagcaggtacttttaaacctt...tgtaaccagctgagctgc 500

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 7442561 seqs, 3452328358 residues

Total number of hits satisfying chosen parameters: 14885122

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	500	100.0	58837	9	US-09-982-091A-5
2	59.8	12.0	337344	19	US-10-388-838-58
3	57.4	11.5	605	13	US-10-027-632-188286
4	57.4	11.5	605	17	US-10-027-632-188286
5	56.2	11.2	43981	17	US-10-292-798-1243
6	55.8	11.2	201	19	US-10-741-601-23975
7	55.8	11.2	201	19	US-10-741-601-23976

C	8	55.8	11.2	201	21	US-10-741-600-64691	Sequence 64691, A
C	9	55.8	11.2	201	21	US-10-741-600-64692	Sequence 64692, A
C	10	55.8	11.2	126872	19	US-10-741-601-5738	Sequence 5738, Ap
C	11	55.8	11.2	126872	21	US-10-741-600-17885	Sequence 17885, A
	12	55.8	11.2	129722	19	US-10-450-826-24	Sequence 24, Appl
	13	55.8	11.2	129722	20	US-10-723-860-844	Sequence 844, App
	14	55.4	11.1	670	22	US-10-915-740A-210	Sequence 210, App
	15	55.4	11.1	57038	22	US-10-287-436A-646	Sequence 646, App
	16	55.4	11.1	92794	22	US-10-287-436A-708	Sequence 708, App
	17	55	11.0	201	19	US-10-741-601-19945	Sequence 19945, A
C	18	55	11.0	201	21	US-10-741-600-55370	Sequence 55370, A
C	19	55	11.0	61177	19	US-10-741-601-5694	Sequence 5694, Ap
C	20	55	11.0	61177	21	US-10-741-600-17768	Sequence 17768, A
C	21	54.8	11.0	27007	9	US-09-764-868-1424	Sequence 1424, Ap
	22	54.8	11.0	105499	19	US-10-322-281-326	Sequence 326, App
	23	54.4	10.9	362	11	US-09-969-034-3487	Sequence 3487, Ap
	24	54.4	10.9	1065	13	US-10-027-632-9474	Sequence 9474, Ap
	25	54.4	10.9	1065	13	US-10-027-632-9475	Sequence 9475, Ap
	26	54.4	10.9	1065	13	US-10-027-632-9476	Sequence 9476, Ap
	27	54.4	10.9	1065	17	US-10-027-632-9474	Sequence 9474, Ap
	28	54.4	10.9	1065	17	US-10-027-632-9475	Sequence 9475, Ap
	29	54.4	10.9	1065	17	US-10-027-632-9476	Sequence 9476, Ap
C	30	54.2	10.8	56510	21	US-10-741-600-17637	Sequence 17637, A
C	31	54.2	10.8	71887	13	US-10-087-192-1654	Sequence 1654, Ap
C	32	54	10.8	176001	17	US-10-210-556-27	Sequence 27, Appl
C	33	54	10.8	186739	17	US-10-210-556-19	Sequence 19, Appl
	34	54	10.8	220224	13	US-10-087-192-1282	Sequence 1282, Ap
	35	53.6	10.7	577	13	US-10-027-632-84682	Sequence 84682, A
	36	53.6	10.7	577	13	US-10-027-632-103818	Sequence 103818, A
	37	53.6	10.7	577	13	US-10-027-632-321553	Sequence 321553, A
	38	53.6	10.7	577	17	US-10-027-632-84682	Sequence 84682, A
	39	53.6	10.7	577	17	US-10-027-632-103818	Sequence 103818, A
C	40	53.6	10.7	577	17	US-10-027-632-321553	Sequence 321553, A
C	41	53.6	10.7	663	13	US-10-027-632-241850	Sequence 241850, A
C	42	53.6	10.7	663	13	US-10-027-632-241851	Sequence 241851, A
C	43	53.6	10.7	663	17	US-10-027-632-241850	Sequence 241850, A
C	44	53.6	10.7	663	17	US-10-027-632-241851	Sequence 241851, A
	45	53.6	10.7	347814	19	US-10-322-696-76	Sequence 76, Appl

#### ALIGNMENTS

RESULT 1  
US-09-982-091A-5  
Sequence 5, Application US/09982091A  
Patent No. US20020151030A1  
GENERAL INFORMATION:  
APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY  
APPLICANT: KUMAGAI, Akiko  
APPLICANT: DUNPHY, William  
TITLE OF INVENTION: CLASPIN PROTEINS AND METHODS OF USE THEREOF  
FILE REFERENCE: CIT1320-1  
CURRENT APPLICATION NUMBER: US/09/982, 091A  
CURRENT FILING DATE: 2002-10-17  
PRIOR APPLICATION NUMBER: US 60/241,246  
PRIOR FILING DATE: 2000-10-17  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5  
LENGTH: 58837  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-982-091A-5

Query Match	Score	DB	Length	Seq					
100.0%	500	9	58837	US-09-982-091A-5					
Best Local Similarity	100.0%	Pred. No.	8.2e-109						
Matches	500	Conservative	0	Mismatches	0	Indels	0	Gaps	0

DB 1 AAGCAGGTACTTTTAACTTTACTCCAGAGGATAGCAGTACTAAATGTACATGCTA 60  
1 AAGCAGGTACTTTTAACTTTACTCCAGAGGATAGCAGTACTAAATGTACATGCTA 60



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 23:18:16 ; Search time 1645.02 Seconds  
(without alignments)  
11569.535 Million cell updates/sec

Title: US-09-982-091A-5\_COPY\_1\_500  
Perfect score: 500  
Sequence: 1 aagcaggtagcttctaact.....tgccaccagcgctgagtcg 500

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_hlc: \*  
4: gb\_est3: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_gsa1: \*  
9: gb\_gsa2: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	138.8	27.8	995	6	CD358047 AGENCOURT
2	58.2	11.6	396	8	AQ436369 HS_5060_B
3	57.8	11.6	796	4	BG548055 602575976
4	55.6	11.1	1101	9	CNS00262 AL097301 Drosophila
5	54.6	10.9	701	5	BX486320 DKFZp686C
6	54.6	10.9	818	8	AQ895736 HS_5439_A
7	54.6	10.9	979	2	BE740888 601593320
8	54.4	10.9	292	1	AI521806 t182f06.x
9	54	10.8	662	9	AG075866 Pan trogl
10	53.8	10.8	317	8	AQ041089 CIT-HSP-2
11	53.8	10.8	507	8	AQ140467 HS_3116_A
12	53.8	10.8	611	5	BQ636112 h04e07.y
13	53.2	10.6	368	1	AI056177 oy57h11.x
14	53.2	10.6	533	5	BX492724 DKFZp781C
15	53.2	10.6	562	1	AU147162 AU147162
16	53.2	10.6	562	8	AQ694806 HS_5486_A
17	53.2	10.6	629	5	BQ020418 UI-H-DP0-
18	53	10.6	368	4	BI055275 PM3-GN037
19	53	10.6	397	4	BI005101 PM3-HN007
20	52.8	10.6	662	9	AG176248 Pan trogl
21	52.8	10.6	542	5	BX484253 DKFZp686P
22	52.6	10.5	436	2	BF872655 IL3-ET011
23	52.6	10.5	665	9	AG157485 Pan trogl
24	52.4	10.5	624	9	AG019812 Homo sapi

C 25	52.2	10.4	552	4	BI850423 imagec 1
C 26	52.2	10.4	698	7	CR771144 DKFZp465H
C 27	52.2	10.4	730	9	AG105601 Pan trogl
C 28	52.2	10.4	776	5	BU853685 AGENCOURT
C 29	52	10.4	228	8	AQ316848 RPI11-96
C 30	52	10.4	425	1	AI669636 tw34G04.x
C 31	52	10.4	852	4	BG429372 602494084
C 32	52	10.4	1101	9	CNS0039G Drosophila
C 33	51.8	10.4	474	9	BF915915 IL3-UT011
C 34	51.8	10.4	721	8	AQ377980 RPI11-15
C 35	51.6	10.3	456	2	BF741873 QV1-HB003
C 36	51.6	10.3	550	7	CR538791 DKFZp459C
C 37	51.6	10.3	853	5	BU529011 AGENCOURT
C 38	51.4	10.3	218	1	AI206863 q130g05.x
C 39	51.4	10.3	353	1	AA372574 EST84511
C 40	51.4	10.3	429	2	AW068007 cn19g10.y
C 41	51.4	10.3	430	2	AW068008 cn19g11.y
C 42	51.4	10.3	437	2	AW504299 UI-HF-BN0
C 43	51.4	10.3	450	2	AW067788 cn19b12.x
C 44	51.4	10.3	456	1	AI336057 q150c03.x
C 45	51.4	10.3	498	8	AQ703721 HS_2160_B

ALIGNMENTS

RESULT 1  
LOCUS CD358047 995 bp mRNA linear EST 29-MAY-2003  
DEFINITION AGENCOURT\_14257551 NIH\_MGC\_180 Homo sapiens CDNA clone  
IMAGE:30387007 5', mRNA sequence.  
ACCESSION CD358047  
VERSION CD358047.1 GI:31129458  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.  
AUTHORS 1 (bases 1 to 995)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein  
CDNA Library Preparation: Invitrogen Corp  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: NDM456 row: j column: 08  
High quality sequence stop: 514.  
Location/Qualifiers  
1. 995  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30387007"  
/lab\_host="DH10B-Ton A ( T1 and T5 phage resistances)"  
/clone\_lib="NIH\_MGC\_180"  
/note="Origin: Testis; Vector: pCMV-SPORT6.1; Site 1: NotI;  
Site 2: EcoRV (destroyed); Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.68 kb. Library was  
constructed by (Invitrogen). Note: this is a NIH\_MGC  
Library."  
ORIGIN  
Query Match 27.8%; Score 138.8; DB 6; Length 995;

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2005, 02:35:27 ; Search time 744.233 Seconds  
(without alignments)  
5019.154 Million cell updates/sec

Title: US-09-982-091A-1\_COPY\_4200\_4740

Perfect score: 541  
Sequence: 1 gtaaatctacctcatagaga.....tcaaaaaaaaaaaaaa 541

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7442561 seqs, 3452328358 residues

Total number of hits satisfying chosen parameters: 14885122

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodaca/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodaca/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodaca/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodaca/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodaca/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodaca/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodaca/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodaca/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodaca/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodaca/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodaca/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodaca/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodaca/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodaca/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodaca/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodaca/1/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodaca/1/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodaca/1/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodaca/1/pubpna/US10G\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodaca/1/pubpna/US10H\_PUBCOMB.seq:\*  
21: /cgn2\_6/ptodaca/1/pubpna/US10I\_PUBCOMB.seq:\*  
22: /cgn2\_6/ptodaca/1/pubpna/US10J\_PUBCOMB.seq:\*  
23: /cgn2\_6/ptodaca/1/pubpna/US11A\_PUBCOMB.seq:\*  
24: /cgn2\_6/ptodaca/1/pubpna/US11\_NEW\_PUB.seq:\*  
25: /cgn2\_6/ptodaca/1/pubpna/US60\_NEW\_PUB.seq:\*  
26: /cgn2\_6/ptodaca/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	541	100.0	4754	9	US-09-982-091A-1 Sequence 1, Appli
2	65.6	12.1	8056	20	US-10-473-126-386 Sequence 386, App
3	65.4	12.1	5979	14	US-10-239-676-18 Sequence 18, Appli
4	65.4	12.1	5979	15	US-10-240-453-26 Sequence 26, Appli
5	64.6	11.9	8305	15	US-10-311-455-1542 Sequence 1542, Ap
6	62.6	11.6	8056	20	US-10-473-126-240 Sequence 240, App
7	62.4	11.5	461	10	US-09-814-353-17724 Sequence 17724, A

C	8	62.2	11.5	8056	20	US-10-473-126-386	Sequence 386, App
	9	61.2	11.3	3673778	16	US-10-312-841-1	Sequence 1, Appli
	10	61	11.3	14919	18	US-10-221-714A-228	Sequence 228, App
C	11	59.8	11.1	3673778	16	US-10-312-841-1	Sequence 1, Appli
	12	59.2	10.9	6101	15	US-10-311-455-1448	Sequence 1448, Ap
	13	59.2	10.9	8056	20	US-10-473-126-240	Sequence 240, App
C	14	58.8	10.9	6101	15	US-10-311-455-1447	Sequence 1447, Ap
	15	58.8	10.9	7851	15	US-10-311-455-1734	Sequence 1734, Ap
C	16	58.6	10.8	6056	15	US-10-311-455-999	Sequence 999, App
	17	58.6	10.8	7167	15	US-10-311-455-373	Sequence 373, App
	18	57.8	10.7	15861	15	US-10-311-455-498	Sequence 498, App
C	19	57.4	10.6	6631	15	US-10-240-453-213	Sequence 213, App
C	20	57.2	10.6	2000	9	US-09-938-842A-3251	Sequence 3251, Ap
C	21	57.2	10.6	2000	11	US-09-938-842A-3251	Sequence 3251, Ap
	22	57.2	10.6	2290	20	US-10-602-494-283	Sequence 283, App
	23	57.2	10.6	17280	18	US-10-221-714A-498	Sequence 498, App
C	24	56.8	10.5	9810	15	US-10-311-455-399	Sequence 399, App
	25	56.2	10.4	38342	18	US-10-221-714A-471	Sequence 471, App
C	26	55.8	10.3	11416	15	US-10-311-455-91	Sequence 91, Appl
	27	55.8	10.3	11416	17	US-10-221-613-19	Sequence 19, Appl
C	28	55.6	10.3	520	19	US-10-021-323-7699	Sequence 7699, Ap
	29	55.6	10.3	11416	15	US-10-311-455-92	Sequence 92, Appl
C	30	55.6	10.3	11416	17	US-10-221-613-20	Sequence 20, Appl
C	31	55.6	10.3	14919	18	US-10-221-714A-228	Sequence 228, App
C	32	55.6	10.3	3673778	16	US-10-312-841-2	Sequence 2, Appli
	33	55.4	10.2	505	19	US-10-437-963-69789	Sequence 69789, A
C	34	55.4	10.2	5875	15	US-10-311-455-262	Sequence 262, App
	35	55.4	10.2	3673778	16	US-10-312-841-2	Sequence 2, Appli
C	36	55	10.2	622	18	US-10-424-599-3380	Sequence 3380, Ap
C	37	54.8	10.1	302	10	US-09-814-353-5335	Sequence 5335, Ap
C	38	54.8	10.1	302	10	US-09-814-353-11622	Sequence 11622, A
	39	54.8	10.1	7851	15	US-10-311-455-1733	Sequence 1733, Ap
	40	54.8	10.1	40324	19	US-10-433-793-179	Sequence 179, App
C	41	54.6	10.1	6794	17	US-10-221-613-134	Sequence 134, App
	42	54.4	10.1	7455	15	US-10-311-455-1732	Sequence 1732, Ap
	43	54.4	10.1	8876	15	US-10-311-455-2050	Sequence 2050, Ap
	44	54.2	10.0	8952	18	US-10-221-714A-167	Sequence 167, App
C	45	54	10.0	1431	9	US-09-774-414-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-982-091A-1  
; Sequence 1, Application US/09982091A  
; Patent No. US20020151030A1  
; GENERAL INFORMATION:  
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY  
; APPLICANT: KUMAGAI, Akiko  
; APPLICANT: DUNPHY, William  
; TITLE OF INVENTION: CLASPIN PROTEINS AND METHODS OF USE THEREOF  
; FILE REFERENCE: CIT1320-1  
; CURRENT APPLICATION NUMBER: US/09/982,091A  
; PRIOR FILING DATE: 2002-10-17  
; PRIOR APPLICATION NUMBER: US 60/241,246  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 4754  
; TYPE: DNA  
; ORGANISM: Xenopus laevis  
US-09-982-091A-1

Query Match 100.0%; Score 541; DB 9; Length 4754;  
Best Local Similarity 100.0%; Pred. No. 4.1e-88;  
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GTAATCTACCTCATAGAGATACAGATACCCCTACAAATAACAGTAATTGTTTACAAC 60  
Db 4200 GTAATCTACCTCATAGAGATACAGATACCCCTACAAATAACAGTAATTGTTTACAAC 4259

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: October 4, 2005, 00:01:51 ; Search time 89.5161 Seconds  
(without alignments)  
9889.010 Million cell updates/sec

Title: US-09-982-091a-1\_COPY\_4200\_4740

Perfect score: 541  
Sequence: 1 gtaatctactctcatagaga.....ctaaaaaataaaaaa 541

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/prodata/1/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58.8	10.9	19124	2	US-08-487-826B-13 Sequence 13, Appl
2	55	10.2	187169	4	US-09-949-016-12776 Sequence 12776, A
3	55	10.2	191569	4	US-09-949-016-15940 Sequence 15940, A
4	54.6	10.1	9447	4	US-09-949-016-14649 Sequence 14649, A
5	54.4	10.1	47781	4	US-09-949-016-16492 Sequence 16492, A
6	54.4	10.1	47781	4	US-09-949-016-16493 Sequence 16493, A
7	54.4	10.1	47781	4	US-09-949-016-16494 Sequence 16494, A
8	54	10.0	1431	3	US-09-316-083-2 Sequence 2, Appl1
9	54	10.0	1431	4	US-09-933-700-2 Sequence 2, Appl1
10	54	10.0	65415	4	US-09-949-016-16699 Sequence 16699, A
11	53.6	9.9	312470	4	US-09-949-016-14043 Sequence 14043, A
12	52.8	9.8	636	3	US-08-998-416-1137 Sequence 1137, Ap
13	52.8	9.8	837	3	US-08-998-416-288 Sequence 288, App
14	52.4	9.7	3350	3	US-08-617-860B-3 Sequence 3, Appl1
15	52.2	9.6	187169	4	US-09-949-016-12776 Sequence 12776, A
16	52.2	9.6	191569	4	US-09-949-016-15940 Sequence 15940, A
17	52	9.6	50383	4	US-09-949-016-17600 Sequence 17600, A
18	51.8	9.6	601	4	US-09-949-016-156535 Sequence 156535,
19	51.8	9.6	615	3	US-08-998-416-186 Sequence 186, App
20	51.8	9.6	231129	4	US-09-949-016-16110 Sequence 16110, A
21	51.8	9.6	266293	4	US-09-949-016-11934 Sequence 11934, A
22	51.6	9.5	640681	4	US-09-790-988-1 Sequence 1, Appl1
23	51	9.4	601	4	US-09-949-016-30530 Sequence 30530, A
24	51	9.4	601	4	US-09-949-016-30531 Sequence 30531, A
25	51	9.4	601	4	US-09-949-016-37149 Sequence 37149, A
26	51	9.4	601	4	US-09-949-016-37150 Sequence 37150, A
27	51	9.4	601	4	US-09-949-016-37163 Sequence 37163, A

C 28	51	9.4	601	4	US-09-949-016-37164	Sequence 37164, A
C 29	51	9.4	601	4	US-09-949-016-145867	Sequence 145867, A
C 30	51	9.4	601	4	US-09-949-016-145868	Sequence 145868, A
C 31	51	9.4	601	4	US-09-949-016-146135	Sequence 146135, A
C 32	51	9.4	601	4	US-09-949-016-146136	Sequence 146136, A
C 33	51	9.4	601	4	US-09-949-016-146403	Sequence 146403, A
C 34	51	9.4	601	4	US-09-949-016-146404	Sequence 146404, A
C 35	51	9.4	205044	4	US-09-949-016-15851	Sequence 15851, A
C 36	51	9.4	205044	4	US-09-949-016-15852	Sequence 15852, A
C 37	51	9.4	205044	4	US-09-949-016-15853	Sequence 15853, A
C 38	51	9.4	223471	4	US-09-949-016-12387	Sequence 12387, A
C 39	51	9.4	223471	4	US-09-949-016-12724	Sequence 12724, A
C 40	51	9.4	223471	4	US-09-949-016-12725	Sequence 12725, A
C 41	50.8	9.4	18989	4	US-09-949-016-13435	Sequence 13435, A
C 42	50.8	9.4	21885	4	US-09-949-016-12600	Sequence 12600, A
C 43	50.8	9.4	101128	4	US-09-949-016-14293	Sequence 14293, A
C 44	50.6	9.4	119153	4	US-09-949-016-12378	Sequence 12378, A
C 45	50.4	9.3	19124	2	US-08-487-826B-13	Sequence 13, Appl1

ALIGNMENTS

RESULT 1  
US-08-487-826B-13  
Sequence 13, Application US/08487826B  
Patent No. 5993827  
GENERAL INFORMATION:  
APPLICANT: Sim, Kim L.  
APPLICANT: Chitnis, Chetan  
APPLICANT: Miller, Louis H.  
APPLICANT: Peterson, David S.  
APPLICANT: Su, Xin-zhaun  
APPLICANT: Wellens, Thomas E.  
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe Martens Olson & Bear  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: California  
COUNTRY: US  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487, 826B  
FILING DATE: 10-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: NIH121.001CPI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19124 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-487-826B-13  
Query Match 10.9%; Score 58.8; DB 2; Length 19124;  
Best Local Similarity 48.8%; Pred. No. 0.0034;



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 23:18:16 ; Search time 1779.91 Seconds  
(without alignments)  
11569.535 Million cell updates/sec

Title: US-09-982-091A-1\_COPY\_4200\_4740  
Perfect score: 541  
Sequence: 1 gtaactctaccatagaga.....ttaaaaaaaaaaaaaaaaaa 541

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gse81:\*  
9: gb\_gse82:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	496.6	91.8	604	4	BJ055644 BJ055644
2	493	91.1	493	2	BE188960 BJ055644/c
3	452	83.5	738	5	BP734671 BP734671
4	365.2	67.5	730	5	BP734936 BP734936
5	299.4	55.3	921	5	BU906840 AGENCOURT
6	160.6	29.7	609	7	CR585851 CR585851
7	159	29.4	568	1	AL879187 AL879187
8	154.2	28.5	366	5	BX780631 BX780631
9	150.6	27.8	737	5	BX783149 BX783149
10	148.4	27.4	550	1	AL881003 AL881003
11	148	27.4	328	1	AL857987 AL857987
12	142	26.2	322	1	AL865252 AL865252
13	133.2	24.6	659	1	AL847560 AL847560
14	78.8	14.6	1101	9	AL069706 Drosophil
15	77.6	14.3	1101	9	AL106896 Drosophil
16	74.6	13.8	1101	9	AL057419 Drosophil
17	73.8	13.6	928	9	CNS00BO1
18	73.4	13.6	1092	9	CNS020K7
19	72.6	13.4	1101	9	CNS00BO1
20	72.2	13.3	1101	9	CNS00EVL
21	70.8	13.1	885	6	CA973527 AGENCOURT
22	70.6	13.0	1101	9	CNS0039G
23	69.6	12.9	1203	9	CNS01603
24	69.4	12.8	1092	9	CNS020K7

25	68.6	12.7	1001	9	CNS0155H	AL105023 Drosophil
26	68.2	12.6	1101	9	CNS00LT2	AL078714 Drosophil
27	68.2	12.6	1225	9	CNS0161D	AL106171 Drosophil
28	67.8	12.5	1101	9	CNS004ZM	AL055440 Drosophil
29	67.6	12.5	1101	9	CNS0021J	AL061936 Drosophil
30	67.6	12.5	1101	9	CNS00EO7	AL069440 Drosophil
31	67	12.4	1200	9	CNS016CO	AL106578 Drosophil
32	66.6	12.3	1200	9	CNS016CO	AL106578 Drosophil
33	66.4	12.3	1190	9	CNS020N7	AL206908 Tetradon
34	65.4	12.1	856	9	AG515880	AG515880 Mus muscu
35	65.4	12.1	869	9	CNS017Z2	AL108680 Drosophil
36	65.4	12.1	935	9	CNS00JKN	AL076398 Drosophil
37	65.4	12.1	1663	3	CR662973	CR662973 Tetradon
38	65.2	12.1	1101	9	CNS00FVE	AL071298 Drosophil
39	65.2	12.1	1780	9	AG320553	AG320553 Mus muscu
40	65	12.0	1101	9	CNS017ZQ	AL108704 Drosophil
41	64.8	12.0	1013	9	CNS06RPQ	AL412260 T7 end of
42	64.2	11.9	660	8	BH183498	BH183498 023 L 07-
43	64.2	11.9	660	9	CNS07ONJ	AL620449 T3 end of
44	63.8	11.8	758	9	AG547036	AG547036 Mus muscu
45	63.8	11.8	926	4	BM358145	BM358145 GA_Ea000

ALIGNMENTS

RESULT 1  
BJ055644/c 604 bp mRNA linear EST 29-SEP-2003  
LOCUS BJ055644 NIBB Mochi normalized Xenopus neurula library Xenopus  
DEFINITION laevis cDNA clone XL050a10 3', mRNA sequence.  
ACCESSION BJ055644  
VERSION BJ055644.1 GI:17468838  
KEYWORDS EST.  
SOURCE Xenopus laevis (African clawed frog)  
ORGANISM Xenopus laevis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
Xenopodinae; Xenopus; Xenopus.

REFERENCE 1 (bases 1 to 604)  
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y.  
TITLE Expressed genes in X. laevis embryo  
JOURNAL Unpublished (2001)  
COMMENT Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp  
The information of this clone is available through the following URL.  
http://xenopus.nibb.ac.jp.

FEATURES  
source location/Qualifiers  
1..604  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="XL050a10"  
/tissue\_type="whole embryo"  
/dev\_stage="stage 15"  
/clone\_lib="NIBB Mochi normalized Xenopus neurula library"

ORIGIN

Query Match 91.8%; Score 496.6; DB 4; Length 604;  
Best Local Similarity 98.8%; Pred.No. 4,3e-78;  
Matches 499; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GTAATCTACCTCATAGAGATACAGATACCTCAAAAATACAGTAATTTGTTTACAAC 60  
DB 512 GTAATTTACCTCATAGAGATACAGATACCTCAAAAATACAGTAATTTGTTTACAAC 453



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 21:10:56 ; Search time 280.519 Seconds  
(without alignments)  
11416.644 Million cell updates/sec

Title: US-09-982-091A-1\_COPY\_4200\_4740

Perfect score: 541

Sequence: 1 gtaaatctacctcatagaga.....ttaaaaaaaaaaaaaaaaaa 541

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn1980a:\*  
2: geneseqn1990a:\*  
3: geneseqn2000a:\*  
4: geneseqn2001a:\*  
5: geneseqn2001b:\*  
6: geneseqn2002a:\*  
7: geneseqn2002b:\*  
8: geneseqn2003a:\*  
9: geneseqn2003b:\*  
10: geneseqn2003c:\*  
11: geneseqn2003d:\*  
12: geneseqn2004a:\*  
13: geneseqn2004b:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	541	100.0	4754	6	ABK52610	Abk52610 DNA encod
2	65.6	12.1	8056	8	ABZ10246	Abz10246 Haematopo
3	65.4	12.1	5979	4	AAS45313	Aas45313 Chemicall
4	65.4	12.1	5979	6	ABK28152	Abk28152 DNA trans
5	64.6	11.9	8305	6	ABL33569	Ab133569 Human imm
6	62.6	11.6	8056	8	ABZ10100	Abz10100 Haematopo
7	62.4	11.5	461	5	ADL43834	Adl43834 Human ova
8	62.2	11.5	8056	8	ABZ10246	Abz10246 Haematopo
9	61	11.3	14919	4	AAS46506	Aas46506 Tumour su
10	59.2	10.9	8056	8	ABL33475	Ab133475 Human imm
11	59.2	10.9	8056	8	ABZ10100	Abz10100 Haematopo
12	58.8	10.9	6101	6	ABL33474	Ab133474 Human imm
13	58.8	10.9	7851	6	ABL33761	Ab133761 Human imm
14	58.8	10.9	19124	2	AAT72882	Aat72882 Plasmodiu
15	58.8	10.9	19124	3	AAS28287	Aas28287 Plasmodiu
16	58.6	10.8	6056	6	ABL33026	Ab133026 Human imm
17	58.6	10.8	7167	6	ABL32400	Ab132400 Human imm
18	58.4	10.8	380	4	AA180354	Aa180354 Human pol
19	57.8	10.7	15861	6	ABL32525	Ab132525 Human imm
20	57.4	10.6	6631	6	ABK28339	Abk28339 DNA trans

C	21	57.2	10.6	2000	6	ABZ15446	Abz15446 Arabidops
C	22	57.2	10.6	17280	4	AAS46772	Aas46772 Tumour su
C	23	56.8	10.5	9810	6	ABL32426	Ab132426 Human imm
C	24	56.2	10.4	38342	4	AAS46745	Aas46745 Tumour su
C	25	56.2	10.4	38342	6	ABK31506	Abk31506 Signal tr
C	26	55.8	10.3	11416	6	ABL32118	Ab132118 Human imm
C	27	55.8	10.3	11416	6	ABL70135	Ab170135 Chemicall
C	28	55.8	10.3	11416	6	AAS61063	Aas61063 Human gen
C	29	55.6	10.3	520	13	ACN52918	Acn52918 Cotton an
C	30	55.6	10.3	1671	2	AAQ24134	Aaq24134 50 KD sub
C	31	55.6	10.3	11416	6	ABL32119	Ab132119 Human imm
C	32	55.6	10.3	11416	6	ABL70136	Ab170136 Chemicall
C	33	55.6	10.3	11416	6	AAS61064	Aas61064 Human gen
C	34	55.6	10.3	14919	4	AAS46506	Aas46506 Tumour su
C	35	55.4	10.2	5875	6	ABL32289	Ab132289 Human imm
C	36	55	10.2	28564	10	ADE61172	Ade61172 Human gen
C	37	55	10.2	28564	10	ADD46006	Add46006 Human gen
C	38	54.8	10.1	302	5	ADL37732	Adl37732 Human ova
C	39	54.8	10.1	302	5	AD172593	Ad172593 Human ova
C	40	54.8	10.1	7851	6	ABL33760	Ab133760 Human imm
C	41	54.8	10.1	40324	6	ABQ67149	Abq67149 Human ang
C	42	54.6	10.1	6794	6	ABK31265	Abk31265 Signal tr
C	43	54.6	10.1	6794	6	ABL70220	Ab170220 Chemicall
C	44	54.6	10.1	6794	6	AAS61175	Aas61175 Human gen
C	45	54.4	10.1	727	4	AA195197	Aa195197 Human neu

ALIGNMENTS

RESULT 1	ABK52610	standard; DNA; 4754 BP.
ID	ABK52610;	
XX		
AC	ABK52610;	
XX		
DT	27-AUG-2002	(first entry)
XX		
DE	DNA encoding Xenopus Claspin protein.	
XX		
KW	Chk1 protein; SQ/TQ motif; isoelectric point; cell cycle progression;	
KW	nuclear localisation signal; DNA replication checkpoint; benign neoplasm;	
KW	cell proliferative disorder; malignant neoplasm; frog; claspin; gene; ds.	
XX		
OS	Xenopus sp.	
XX		
FH	Key	location/Qualifiers
FT	CDS	71..3928
FT		/*tag= a
FT		/product= "Xenopus Claspin protein"
XX		
PN	W0200233115-A2.	
XX		
PD	25-APR-2002.	
XX		
PF	17-OCT-2001; 2001WO-US032316.	
XX		
PR	17-OCT-2000; 2000US-0241246P.	
XX		
PA	(CALY ) CALIFORNIA INST OF TECHNOLOGY.	
XX		
PI	Kumagai A, Dunphy WG;	
XX		
DR	WPI; 2002-454610/48.	
DR	P-PSDB; AAU97586.	
XX		
PT	Novel Claspin polypeptide specifically interacting with chk1 protein	
PT	useful for identifying compound that modulates cell cycle progression and	
PT	for treating cell proliferative disorder like neoplasm.	
XX		
PS	Claim 7; Fig 1; 97pp; English.	
XX		
CC	The present invention relates to a new substantially pure Claspin	

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 21:21:46 ; Search time 1857.72 Seconds  
(without alignments)  
14110.995 Million cell updates/sec

Title: US-09-982-091a-1\_COPY\_4200\_4740  
Perfect score: 541  
Sequence: 1 gtaatctactcctcatagaga.....ttaaaaaaaaaaaaaa 541

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_ey:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	541	100.0	4754	5	AF297867	AF297867 Xenopus l
2	73.2	13.5	253001	3	AE014834	AE014834 Plasmodi
3	69.4	12.8	110000	2	PFMAL13_02	Continuation (3 of
4	68.6	12.7	254436	3	AE014827	AE014827 Plasmodi
5	68.4	12.6	201254	2	CR391938	CR391938 Danio rer
6	67.8	12.5	58891	2	AC021479	AC021479 Mus muscu
7	67.8	12.5	254050	3	PFA929358	AL929358 Plasmodi
8	67.2	12.4	77835	2	PFMAL13P2_3	Continuation (4 of
9	67.2	12.4	340552	3	PFA929354	AL929354 Plasmodi
10	66.8	12.3	148484	2	CR450693	CR450693 Danio rer
11	66.6	12.3	242140	2	CR456633	CR456633 Danio rer
12	66.4	12.3	159618	9	AC096750	AC096750 Homo sapi
13	66.4	12.3	179066	2	CR450782	CR450782 Danio rer
14	66.2	12.2	110000	2	PFMAL13_15	Continuation (16 o
15	66.2	12.2	199475	2	CR318662	CR318662 Danio rer
16	66	12.2	161444	2	BX897740	BX897740 Danio rer
17	66	12.2	185596	9	AC021553	AC021553 Homo sapi
18	66	12.2	250029	3	AE014816	AE014816 Plasmodi
19	65.8	12.2	64707	3	AC115607	AC115607 Dictyoste

20	65.6	12.1	8056	6	AX599046	AX599046 Sequence
21	65.6	12.1	27694	8	HASMT	D31785 Pichia cana
22	65.6	12.1	141946	2	CR396585	CR396585 Danio rer
23	65.4	12.1	5979	6	AX277855	AX277855 Sequence
24	65.4	12.1	5979	6	AX323538	AX323538 Sequence
25	65.4	12.1	152209	9	HS1108D11	AL034419 Human DNA
26	65.2	12.1	54441	3	AC115584	AC115584 Dictyoste
27	65.2	12.1	254050	3	PFA929358	AL929358 Plasmodi
28	65	12.0	64707	3	AC115607	AC115607 Dictyoste
29	65	12.0	110000	2	PFMAL13_23	Continuation (24 o
30	65	12.0	183373	2	CR381594	CR381594 Danio rer
31	65	12.0	204652	2	PFMAL13P6	AL049183 Plasmodi
32	64.8	12.0	110000	3	AC116957_0	AC116957 Dictyoste
33	64.8	12.0	191887	2	CR407563	CR407563 Danio rer
34	64.8	12.0	249995	3	AE014840	AE014840 Plasmodi
35	64.8	12.0	299050	1	AP004173	AP004173 Mycoplasma
36	64.6	11.9	8305	6	AX346471	AX346471 Sequence
37	64.6	11.9	110000	3	PFMAL1P2_1	Continuation (2 of
38	64.6	11.9	207198	2	CR376829	CR376829 Danio rer
39	64.6	11.9	242140	2	CR456633	CR456633 Danio rer
40	64.6	11.9	330199	2	BX901880	BX901880 Danio rer
41	64.4	11.9	153477	2	AC006278	AC006278 Plasmodi
42	64.4	11.9	154765	2	CR847899	CR847899 Danio rer
43	64.4	11.9	183614	2	CR847995	CR847995 Danio rer
44	64.4	11.9	218083	2	AC142177	AC142177 Rattus no
45	64.4	11.9	251551	3	AE014844	AE014844 Plasmodi

ALIGNMENTS

RESULT 1	AF297867	4754 bp	mRNA	linear	VRT 28-MAY-2002
LOCUS	AF297867	Xenopus laevis Claspin mRNA, complete cds.			
DEFINITION	AF297867				
ACCESSION	AF297867				
VERSION	AF297867.2	GI:21233661			
KEYWORDS					
SOURCE					
ORGANISM	Xenopus laevis (African clawed frog)				
REFERENCE	1 (bases 1 to 4754)				
AUTHORS	Kumagai, A. and Dunphy, W.G.				
TITLE	Claspin, a novel protein required for the activation of Chk1 during a DNA replication checkpoint response in Xenopus egg extracts				
JOURNAL	Mol. Cell 6 (4), 839-849 (2000)				
MEDLINE	21000493				
PUBMED	11090622				
REFERENCE	2 (bases 1 to 4754)				
AUTHORS	Kumagai, A. and Dunphy, W.G.				
TITLE	Direct Submission				
JOURNAL	Submitted (21-AUG-2000) Biology, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA				
REFERENCE	3 (bases 1 to 4754)				
AUTHORS	Kumagai, A. and Dunphy, W.G.				
TITLE	Direct Submission				
JOURNAL	Submitted (28-MAY-2002) Biology, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA				
REMARK	Sequence update by submitter				
COMMENT	On May 28, 2002 this sequence version replaced gi:10944335.				
FEATURES	Location/Qualifiers				
source	1..4754				
	/organism="Xenopus laevis"				
	/mol_type="mRNA"				
	/db_xref="taxon:8355"				
	71..3928				
	/note="required for the activation of Chk1"				
	/codon_start=1				
	/product="Claspin"				
	/protein_id="AAC24516.2"				
	/db_xref="GI:21233662"				

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2005, 02:35:27 ; Search time 687.831 Seconds  
(without alignments)  
5019.154 Million cell updates/sec

Title: US-09-982-091A-1\_COPY\_1\_500

Perfect score: 500  
Sequence: 1 acgcataggcgcgcaattcc.....agaacatgatagcttgagaga 500

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7442561 seqs, 3452328358 residues

Total number of hits satisfying chosen parameters: 14885122

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*  
22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	500	100.0	4754	9	US-09-982-091A-1
2	47.6	9.5	2859	19	US-10-437-963-97519
3	46.2	9.2	583	9	US-09-864-761-20772
4	46.2	9.2	1959	9	US-09-864-761-4012
5	46.2	9.2	114771	20	US-10-723-860-458
6	44.4	8.9	2717	19	US-10-451-467A-237
7	44.2	8.8	421	10	US-09-803-719-1138

8	44.2	8.8	496	10	US-09-918-995-21694	Sequence 21694, A
9	44.2	8.8	4756	9	US-09-982-091A-3	Sequence 3, Appli
10	44.2	8.8	4804	20	US-10-357-930-30217	Sequence 30217, A
11	44	8.8	393	10	US-09-803-719-853	Sequence 853, App
12	44	8.8	401	20	US-10-357-930-15892	Sequence 15892, A
13	44	8.8	422	10	US-09-803-719-1139	Sequence 1139, Ap
14	44	8.8	452	20	US-10-357-930-45720	Sequence 45720, A
15	43.8	8.8	387	19	US-10-021-323-941	Sequence 941, App
16	43.6	8.7	344	19	US-10-437-963-5076	Sequence 5076, Ap
17	43.2	8.6	766	9	US-09-864-761-19608	Sequence 19608, A
18	43.2	8.6	1944	9	US-09-864-761-2825	Sequence 2825, Ap
19	42.6	8.5	408	21	US-10-275-323A-13	Sequence 13, Appl
20	42.6	8.5	519	19	US-10-437-963-101926	Sequence 101926,
21	42	8.4	408	10	US-09-803-719-777	Sequence 777, App
22	42	8.4	1452	17	US-10-158-057-65	Sequence 65, Appl
23	41.8	8.4	58837	9	US-09-982-091A-5	Sequence 5, Appli
24	41.4	8.3	1083	22	US-10-450-763-12549	Sequence 12549, A
25	41.4	8.3	58985	10	US-09-901-152-3	Sequence 3, Appli
26	41.4	8.3	58985	21	US-10-483-593-3	Sequence 3, Appli
27	41.4	8.3	143601	10	US-09-855-824-3	Sequence 3, Appli
28	41.4	8.3	143601	20	US-10-476-542-3	Sequence 3, Appli
29	41.2	8.2	327	9	US-09-864-761-28059	Sequence 28059, A
30	41.2	8.2	3585	19	US-10-437-963-62522	Sequence 62522, A
31	41.2	8.2	63693	19	US-10-741-601-5650	Sequence 5650, Ap
32	41	8.2	1284	16	US-10-032-585-6386	Sequence 6386, Ap
33	40.6	8.1	872	9	US-09-925-297-358	Sequence 358, App
34	40.6	8.1	1984	14	US-10-050-704-70	Sequence 70, Appl
35	40.6	8.1	1984	19	US-10-798-512-70	Sequence 70, Appl
36	40.6	8.1	4623	9	US-09-864-864-288	Sequence 288, App
37	40.6	8.1	4633	22	US-10-450-763-5717	Sequence 5717, Ap
38	40.6	8.1	4637	18	US-10-363-616-11	Sequence 11, Appl
39	40.6	8.1	4684	18	US-10-363-616-12	Sequence 12, Appl
40	40.6	8.1	4689	22	US-10-450-763-9626	Sequence 9626, Ap
41	40.6	8.1	5136	22	US-10-450-763-20519	Sequence 20519, A
42	40.4	8.1	456	9	US-09-864-761-11468	Sequence 11468, A
43	40.4	8.1	591	19	US-10-021-323-13554	Sequence 13554, A
44	40.4	8.1	671	14	US-10-184-644-346	Sequence 346, App
45	40.4	8.1	671	14	US-10-184-634-346	Sequence 346, App

ALIGNMENTS

RESULT 1

US-09-982-091A-1

; Sequence 1, Application US/09982091A

; Patent No. US20020151030A1

; GENERAL INFORMATION:

; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY

; APPLICANT: KUMAGAI, Akiko

; APPLICANT: DUNPHY, William

; TITLE OF INVENTION: CLASPIN PROTEINS AND METHODS OF USE THEREOF

; FILE REFERENCE: CIT1320-1

; CURRENT APPLICATION NUMBER: US/09/982, 091A

; CURRENT FILING DATE: 2002-10-17

; PRIOR APPLICATION NUMBER: US 60/241,246

; PRIOR FILING DATE: 2000-10-17

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 4754

; TYPE: DNA

; ORGANISM: Xenopus laevis

US-09-982-091A-1

Query Match 100.0%; Score 500; DB 9; Length 4754;  
Best Local Similarity 100.0%; Pred. No. 9.7e-118;  
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACGCATAGGCGCGAATTCACAGCGCGAGTAGTGAGATTGCGGAGCTGTCAACACC 60  
Db 1 ACGCATAGGCGCGAATTCACAGCGCGAGTAGTGAGATTGCGGAGCTGTCAACACC 60

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 4, 2005, 00:01:51 ; Search time 82.7321 Seconds  
(without alignments)  
9889.010 Million cell updates/sec

Title: US-09-982-091a-1\_COPY\_1\_500  
Perfect score: 500  
Sequence: 1 acgcacataggcgccgaattcc.....agacacatgatagcttgaga 500

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	70.6	14.1	7218	1 US-08-232-463-14	Sequence 14, Appl
2	44.2	8.8	4756	4 US-09-949-016-4455	Sequence 4455, Ap
3	43.8	8.8	767677	4 US-09-949-016-12147	Sequence 12147, A
4	43.8	8.8	767677	4 US-09-949-016-17361	Sequence 17361, A
5	43	8.6	832	4 US-09-621-976-2813	Sequence 2813, Ap
6	41.8	8.4	53737	4 US-09-949-016-16197	Sequence 16197, A
7	40.6	8.1	642	4 US-09-248-796A-4755	Sequence 4755, Ap
8	40.2	8.0	1832	4 US-09-976-594-537	Sequence 537, App
9	40.2	8.0	1961	4 US-09-976-594-535	Sequence 535, App
10	40.2	8.0	10304	4 US-09-627-465B-1	Sequence 1, Appli
C 11	39.8	8.0	278866	4 US-09-949-016-13922	Sequence 13922, A
C 12	39.8	8.0	278866	4 US-09-949-016-13923	Sequence 13923, A
C 13	39.8	8.0	278866	4 US-09-949-016-13924	Sequence 13924, A
C 14	39.8	8.0	278866	4 US-09-949-016-13925	Sequence 13925, A
C 15	39.8	8.0	278866	4 US-09-949-016-13926	Sequence 13926, A
C 16	39.8	8.0	278866	4 US-09-949-016-14699	Sequence 14699, A
C 17	39.8	8.0	278866	4 US-09-949-016-14700	Sequence 14700, A
C 18	39.8	8.0	278866	4 US-09-949-016-14701	Sequence 14701, A
C 19	39.8	8.0	278866	4 US-09-949-016-14702	Sequence 14702, A
C 20	39.8	8.0	278866	4 US-09-949-016-14703	Sequence 14703, A
21	39.6	7.9	3561	4 US-09-248-796A-2080	Sequence 2080, Ap
22	39.6	7.9	31391	4 US-09-949-016-14319	Sequence 14319, A
23	39	7.8	289	3 US-09-007-005-17	Sequence 17, Appl
24	39	7.8	289	3 US-09-244-796-17	Sequence 17, Appl
25	39	7.8	2130	4 US-09-248-796A-4942	Sequence 4942, Ap
26	38.8	7.8	1002	4 US-09-248-796A-1249	Sequence 1249, Ap
27	38.6	7.7	1566	4 US-09-248-796A-4899	Sequence 4899, Ap

28	38.6	7.7	192302	4 US-09-949-016-15270	Sequence 15270, A
29	38.4	7.7	834	4 US-09-248-796A-1248	Sequence 1248, Ap
30	38.2	7.6	4344	4 US-09-601-198-165	Sequence 165, App
31	38	7.6	1284	4 US-09-107-532A-825	Sequence 825, App
32	38	7.6	1596	4 US-09-248-796A-6132	Sequence 6132, Ap
33	38	7.6	3114	3 US-09-134-001C-1957	Sequence 1957, Ap
34	37.8	7.6	906	4 US-09-710-279-3175	Sequence 3175, Ap
35	37.8	7.6	1239	4 US-09-286-981B-22	Sequence 22, Appl
36	37.8	7.6	1620	4 US-09-134-001C-1868	Sequence 1868, Ap
37	37.8	7.6	1623	3 US-09-134-001C-1868	Sequence 22, Appl
38	37.8	7.6	2295	1 US-08-375-300-3	Sequence 3, Appli
39	37.8	7.6	2295	3 US-09-177-431-3	Sequence 3, Appli
40	37.8	7.6	2295	5 PCT-US95-16930-3	Sequence 3, Appli
41	37.8	7.6	3185	4 US-09-710-279-3758	Sequence 3758, Ap
C 42	37.8	7.6	3209	4 US-09-710-279-3494	Sequence 3494, Ap
C 43	37.8	7.6	3885	4 US-09-710-279-3883	Sequence 3883, Ap
44	37.8	7.6	4028	4 US-09-710-279-4316	Sequence 4316, Ap
45	37.8	7.6	4080	1 US-08-375-300-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-08-232-463-14/c  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 INMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: PTZgpt-Fls  
US-08-232-463-14  
Query Match 14.1%; Score 70.6; DB 1; Length 7218;

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 23:18:16 (Search time 1645.02 Seconds  
(without alignments)  
11569.535 Million cell updates/sec

Title: US-09-982-091a-1\_COPY\_1\_500  
Perfect score: 500  
Sequence: 1 acgcacagggcggaattcc.....agaacatgatagcttgaga 500

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hnc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	459.2	91.8	981	6	CA974658 AGENCOURT
2	447.4	89.5	915	6	CA791641 AGENCOURT
3	432.8	86.6	926	5	BU907710 AGENCOURT
4	400.8	80.2	955	6	CA982465 AGENCOURT
5	286.4	57.3	661	1	AL870060 AGENCOURT
6	282.6	56.5	858	1	AL863727 AGENCOURT
7	278.6	55.7	571	7	CF783399 AGENCOURT
8	278.6	55.7	577	7	CF783398 AGENCOURT
9	273.6	54.7	543	1	AL896540 AGENCOURT
10	273.6	54.7	561	5	BX774034 AGENCOURT
11	273.6	54.7	643	1	AL957528 AGENCOURT
12	246	49.2	631	1	AL850776 AGENCOURT
13	243.4	48.7	649	1	AL863524 AGENCOURT
14	242.8	48.6	634	1	AL873704 AGENCOURT
15	242.8	48.6	656	1	AL866230 AGENCOURT
16	200.6	40.1	933	5	BX756585 AGENCOURT
17	182.4	36.5	663	1	AL891701 AGENCOURT
18	170.8	34.1	553	1	AL866562 AGENCOURT
19	138.4	27.7	494	1	AL874978 AGENCOURT
20	124.6	24.9	608	1	AL881027 AGENCOURT
21	97	15.4	537	1	AL882988 AGENCOURT
22	76.8	11.4	485	1	AL889102 AGENCOURT
23	57	11.4	1101	9	CNS0039C

C 25	56.2	11.2	849	8	AZ546009	AZ546009	ENTFW53TF
C 26	55.2	11.0	648	2	BE349803	BE349803	hq43c11.x
C 27	53.6	10.7	905	8	AZ550256	AZ550256	ENTBVS8TR
C 28	53	10.6	629	2	BE329068	BE329068	hq36d07.x
C 29	52.6	10.5	509	5	BQ599603	BQ599603	MI-P-E6-a
C 30	52.6	10.5	630	5	BX665807	BX665807	BX665807
C 31	52.6	10.5	630	5	BX668089	BX668089	BX668089
C 32	52.6	10.5	869	5	BP166258	BP166258	BP166258
C 33	52.4	10.5	843	8	AZ551618	AZ551618	ENTDV54TR
C 34	52.4	10.5	908	8	AZ548467	AZ548467	ENTEX30TR
C 35	52.4	10.5	931	8	BH160272	BH160272	ENTOV49TR
C 36	52	10.4	640	2	BE349846	BE349846	hq43h07.x
C 37	51.8	10.4	656	2	BE349729	BE349729	hq42c05.x
C 38	51.4	10.3	976	8	BH149983	BH149983	ENTQD93TF
C 39	51.4	10.3	1490	8	CC293290	CC293290	CH261-62J
C 40	50.8	10.2	888	8	AZ528430	AZ528430	ENTCO24TR
C 41	48.6	9.7	877	8	AZ531291	AZ531291	ENTBQ34TR
C 42	48.6	9.7	912	8	AZ551092	AZ551092	ENTFJ22TF
C 43	48.4	9.7	507	5	BQ739620	BQ739620	PFEStoab4
C 44	48.4	9.7	836	9	BX143478	BX143478	Danio rer
C 45	48.4	9.7	900	8	AZ549980	AZ549980	ENTDD94TF

ALIGNMENTS

RESULT 1  
CA974658  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CA974658 981 bp mRNA linear EST 27-FEB-2003  
AGENCOURT 10748008 Wellcome CRC psk egg Xenopus laevis cDNA clone  
IMAGE:6327325.5', mRNA sequence.  
CA974658.1 GI:27507312  
EST.  
Xenopus laevis (African clawed frog)  
Xenopus laevis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
Xenopodinae; Xenopus; Xenopus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: N. Garrett, P. Lemaire, A.M. Zorn, and J.B. Gurdon (Wellcome/CRC Institute)

cDNA Library Preparation: N. Garrett, P. Lemaire, A.M. Zorn, and J.B. Gurdon (Wellcome/CRC Institute)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: XGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: L1AM13776 row: 1 column: 14  
High quality sequence start: 17  
High quality sequence stop: 724.  
location/Qualifiers

FEATURES  
source

1. 981  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="IMAGE:6327325"  
/tissue\_type="egg"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="Wellcome CRC psk egg"  
/note="Vector: pBluescript SK-; Site\_1: NotI; Site\_2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Library was constructed by N. Garrett, P. Lemaire, A.M. Zorn, and J.B. Gurdon (Wellcome/CRC Institute)."  
Note: This is a Xenopus Gene Collection (XGC) library."

ORIGIN

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 21:10:56 ; Search time 259.259 Seconds  
(without alignments)  
11416.644 Million cell updates/sec

Title: US-09-982-091A-1\_COPY\_1\_500  
Perfect score: 500  
Sequence: 1 acgcataggcgcgcaattcc.....agaacatgatgacttgaga 500

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	500	100.0	4754	6	ABK52610	AbK52610 DNA encod
2	48.4	9.7	1998	3	AAA70212	Aaa70212 Plasmodi
3	46.2	9.2	583	4	AAI23356	Aai23356 Probe #13
4	46.2	9.2	583	4	ABA68463	AbA68463 Human foe
5	46.2	9.2	583	4	AAI48680	Aai48680 Probe #17
6	46.2	9.2	583	4	ABA50512	AbA50512 Human bre
7	46.2	9.2	583	4	ABA35452	AbA35452 Probe #13
8	46.2	9.2	583	4	AAK42604	AaK42604 Human bon
9	46.2	9.2	583	4	AAK16834	AaK16834 Human bra
10	46.2	9.2	583	4	ABS42222	AbS42222 Human liv
11	46.2	9.2	583	5	AAI08992	Aai08992 Probe #89
12	46.2	9.2	583	6	ABS16657	AbS16657 Human gen
13	46.2	9.2	1959	4	AAI14146	Aai14146 Probe #40
14	46.2	9.2	1959	4	ABA55872	AbA55872 Human foe
15	46.2	9.2	1959	4	AAI35531	Aai35531 Probe #42
16	46.2	9.2	1959	4	ABA45382	AbA45382 Human bre
17	46.2	9.2	1959	4	ABA25546	AbA25546 Probe #40
18	46.2	9.2	1959	4	AAK29571	AaK29571 Human bon
19	46.2	9.2	1959	4	AAK04090	AaK04090 Human bra
20	46.2	9.2	1959	4	ABS29202	AbS29202 Human liv

ALIGNMENTS

21	46.2	9.2	1959	5	AAI03994	Aai03994 Probe #39
22	46.2	9.2	1959	6	ABS04127	AbS04127 Human gen
23	46.2	9.2	114771	12	ADQ17641	Adq17641 Human sof
24	44.4	8.9	2717	6	ABQ76406	AbQ76406 S. cerevi
25	44.4	8.9	231004	12	ADQ97855	Adq97855 Mouse can
26	44.2	8.8	421	4	AAS38080	Aas38080 Novel hum
27	44.2	8.8	496	9	ACH34482	Ach34482 Human end
28	44.2	8.8	4414	10	ADC30694	Adc30694 Human nov
29	44.2	8.8	4755	6	ABK52611	AbK52611 DNA encod
30	44.2	8.8	4804	5	ABV30199	Abv30199 Human pro
31	44	8.8	393	4	AAS37795	Aas37795 Novel hum
32	44	8.8	401	5	ABV15901	Abv15901 Human pro
33	44	8.8	406	5	AAF66639	Aaf66639 Novel hum
34	44	8.8	411	5	AAF66638	Aaf66638 Novel hum
35	44	8.8	422	4	AAS38081	Aas38081 Novel hum
36	44	8.8	452	5	ABV45701	Abv45701 Human pro
37	43.8	8.8	387	13	ACN46160	Acn46160 Cotton pr
38	43.2	8.6	766	4	AAI22114	Aai22114 Probe #12
39	43.2	8.6	766	4	ABA67193	AbA67193 Human foe
40	43.2	8.6	766	4	AAI47409	Aai47409 Probe #16
41	43.2	8.6	766	4	ABA49279	AbA49279 Human bre
42	43.2	8.6	766	4	ABA34288	AbA34288 Probe #12
43	43.2	8.6	766	5	AAI07812	Aai07812 Probe #78
44	43.2	8.6	1299	3	AAA70111	Aaa70111 Plasmodi
45	43.2	8.6	1944	4	AAI12874	Aai12874 Probe #28

RESULT 1

ABK52610

ID ABK52610 standard; DNA; 4754 BP.

XX

AC ABK52610;

XX

DT 27-AUG-2002 (first entry)

XX

DE DNA encoding Xenopus Claspin protein.

XX

KW Chk1 protein; SQ/TQ motif; isoelectric point; cell cycle progression;

KW nuclear localisation signal; DNA replication checkpoint; benign neoplasm;

KW cell proliferative disorder; malignant neoplasm; frog; claspin; gene; ds.

XX

OS Xenopus sp.

XX

FH

FH Key

FT CDS

FT 71..3928

FT /\*tag= a

FT /product= "Xenopus Claspin protein"

PN

XX

WO200233115-A2.

XX

PD 25-APR-2002.

XX

PF 17-OCT-2001; 2001WO-US032316.

XX

PR 17-OCT-2000; 2000US-0241246P.

XX

PA (CALY ) CALIFORNIA INST OF TECHNOLOGY.

XX

PI Kumagai A, Dunphy WG;

XX

DR WPI; 2002-454610/48.

DR P-PSDB; AAU97586.

XX

PT Novel Claspin polypeptide specifically interacting with chk1 protein

PT useful for identifying compound that modulates cell cycle progression and

PT for treating cell proliferative disorder like neoplasm.

XX

PS Claim 7; Fig 1; 97pp; English.

XX

CC The present invention relates to a new substantially pure Claspin



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 21:21:46 ; Search time 1716.93 Seconds  
(without alignments)  
14110.995 Million cell updates/sec

Title: US-09-982-091A-1\_COPY\_1\_500

Perfect score: 500  
Sequence: 1 acgcacataggcgcgcaattcc.....agaacatgatagcttgagaga 500

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_ba:\*\n2: gb\_hcg:\*\n3: gb\_in:\*\n4: gb\_om:\*\n5: gb\_ov:\*\n6: gb\_pat:\*\n7: gb\_ph:\*\n8: gb\_pl:\*\n9: gb\_pr:\*\n10: gb\_ro:\*\n11: gb\_sts:\*\n12: gb\_sy:\*\n13: gb\_un:\*\n14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	500	100.0	4754	5	AF297867 Xenopus t
2	273.6	54.7	1044	5	CR761574 Xenopus t
3	70.6	14.1	7218	6	I66494 Sequence 14
4	64.8	13.0	110000	2	PFMAL13_24
5	56.4	11.3	164399	3	PFMAL13P6
6	54.8	11.0	151341	5	AL929536
7	54.8	11.0	158683	2	BX927349
8	54.4	10.9	794	5	BX935928
9	53.6	10.7	1211	10	RATSBP
10	53.2	10.6	143342	2	CR388191
11	53.2	10.6	347582	3	PFMAL4P1
12	52	10.4	110000	2	PFMAL13_07
13	52	10.4	250589	2	AC130139
14	51.6	10.3	347050	3	PPA929351
15	51.4	10.3	110000	2	PFMAL7P1_09
16	51.4	10.3	205429	2	AC005506
17	51.4	10.3	253132	3	AE014846
18	51	10.2	110000	2	PFMAL7P1_11
19	51	10.2	110000	2	PFMAL8P1_02

20	50.8	10.2	61052	2	AC123513	AC123513 Dictyoste
21	50.8	10.2	115758	9	AC104634	AC104634 Homo sapi
22	50.8	10.2	136240	3	AC117070	AC117070 Dictyoste
23	50.4	10.1	250823	3	AE014821	AE014821 Plasmodiu
24	50.4	10.1	251448	3	AE014819	AE014819 Plasmodiu
25	50.4	10.1	254050	3	PPA929358	AL929358 Plasmodiu
26	50	10.0	158548	3	PFMAL3P2	AL034558 Plasmodiu
27	50	10.0	250029	3	AE014838	AE014838 Plasmodiu
28	49.8	10.0	34347	2	BX957360	BX957360 Danio rer
29	49.8	10.0	124801	2	AC149906	AC149906 Strongylo
30	49.8	10.0	199698	2	BX004991	BX004991 Danio rer
31	49.8	10.0	215313	9	AC100791	AC100791 Homo sapi
32	49.4	9.9	72383	2	AC027746	AC027746 Homo sapi
33	49.4	9.9	164347	9	AC104805	AC104805 Homo sapi
34	49.2	9.8	974	3	PPARP3EX	Y08925 P.falciparu
35	48.8	9.8	418	9	BC062215	BC062215 Homo sapi
36	48.8	9.8	115990	5	BX890614	BX890614 Zebrafish
37	48.4	9.7	3016	3	PPASANTA	M18824 P.faciaparum
38	48.4	9.7	16010	3	AE001410	AE001410 Plasmodiu
39	48.4	9.7	175127	2	AC016048	AC016048 Homo sapi
40	48.4	9.7	186147	2	CR626891	CR626891 Danio rer
41	48.4	9.7	194969	2	CR847853	CR847853 Danio rer
42	48.4	9.7	250022	3	AE014824	AE014824 Plasmodiu
43	48.2	9.6	96249	9	AC116612	AC116612 Homo sapi
44	47.8	9.6	94534	5	AL929250	AL929250 Zebrafish
45	47.8	9.6	313050	3	PPA929352	AL929352 Plasmodiu

ALIGNMENTS

RESULT 1	AF297867	4754 bp	mRNA	linear	VRT 28-MAY-2002
LOCUS	AF297867				
DEFINITION	Xenopus laevis Claspin mRNA, complete cds.				
ACCESSION	AF297867				
VERSION	AF297867.2	GI:21233661			
KEYWORDS					
SOURCE					
ORGANISM	Xenopus laevis (African clawed frog)				
REFERENCE					
AUTHORS	Kumagai, A. and Dunphy, W.G.				
TITLE	1 (bases 1 to 4754)				
JOURNAL	Mol. Cell 6 (4), 839-849 (2000)				
PUBMED	21000493				
REFERENCE					
AUTHORS	Kumagai, A. and Dunphy, W.G.				
TITLE	Direct Submision				
JOURNAL	Submitted (21-AUG-2000) Biology, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA				
REFERENCE					
AUTHORS	Kumagai, A. and Dunphy, W.G.				
TITLE	3 (bases 1 to 4754)				
JOURNAL	Submitted (28-MAY-2002) Biology, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA				
REMARK	Sequence update by submitter				
COMMENT	On May 28, 2002 this sequence version replaced gi:10944335.				
FEATURES					
source	Location/Qualifiers				
	1..4754				
	/organism="Xenopus laevis"				
	/mol_type="mRNA"				
	/db_xref="taxon:8355"				
	71..3928				
	/note="required for the activation of Chk1"				
	/codon_start=1				
	/product="Claspin"				
	/protein_id="AA024516.2"				
	/db_xref="GI:21233662"				